
 W E S E R E
 (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MParch_PP protein - protein database search, using Smith-Waterman algorithm
 Run on: Sat May 13 09:23:13 2000; MasPar time 3.41 Seconds
 194.564 Million cell updates/sec
 Tabular output not generated.

Title: >US-09-331-631-28
 Description: (1-28) from US09331631.pep
 Perfect Score: 225
 Sequence: 1 LDPTRQQLCQMRCCQDEKDPKQOQCK 28

Scoring table: PAM 150
 Gap 15

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: a:geneseq35
 1:geneseqp

Statistics: Mean 20.583; Variance 75.953; scale 0.271

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	225	100.0	28	1 W62841	Stenocarpus sinuatus a	1.69e-15
2	119	52.0	625	1 W62830	Macadamia integrifolia	2.93e-04
3	117	52.0	525	1 W62831	Theobroma cacao antimi	4.68e-04
4	117	52.0	566	1 R20181	Sequence encoded by 67	4.68e-04
5	113	50.2	666	1 W62828	Macadamia integrifolia	1.19e-03
6	111	49.3	666	1 W62829	Macadamia integrifolia	1.88e-03
7	104	46.2	805	1 R80561	Murine Ah receptor pro	5.79e-02
8	96	42.7	590	1 W62832	Gossypium hirsutum ant	1.61e-00
9	81	36.0	918	1 R12223	Human androgen recepto	1.61e-00
10	81	36.0	919	1 W14783	Androgen receptor.	1.61e-00
11	81	36.0	919	1 P93109	Human androgen recepto	1.61e-00
12	80	35.6	539	1 W33628	Yeast transcriptional	2.00e+00
13	76	33.8	1420	1 W81025	AIb1 (Amplified in bre	4.73e+00
14	74	32.9	351	1 R31599	Chicken nov protein.	7.25e+00
15	74	32.9	1313	1 W60213	Spinocerebellar ataxia	7.25e+00
16	73	32.4	760	1 W29450	Programmed cell death	8.96e+00
17	72	32.0	86	1 W95078	GST-HD fusion protein	1.11e+01
18	72	32.0	86	1 W95073	GST-HD fusion protein	1.11e+01
19	72	32.0	94	1 W95080	GST-HD fusion protein	1.11e+01
20	72	32.0	94	1 W95075	GST-HD fusion protein	1.11e+01
21	72	32.0	108	1 W95076	Amino acid sequence of	1.11e+01
22	72	32.0	108	1 W95071	Amino acid sequence of	1.11e+01
23	72	32.0	171	1 W99022	Human huntingtin prote	1.11e+01

24	72	32.0	371	1 W73369	Epitope tagged TBP pro	1.11e+01
25	72	32.0	731	1 W96312	Human small conductanc	1.11e+01
26	72	32.0	736	1 W63717	Human hsk3 protein.	1.11e+01
27	72	32.0	3144	1 R58777	Protein encoded by Hun	1.11e+01
28	72	32.0	3144	1 W44742	Human huntingtin prote	1.11e+01
29	72	32.0	3144	1 W09871	Human huntingtin.	1.11e+01
30	72	32.0	3144	1 W36887	Previously undescribed	1.11e+01
31	70	31.1	919	1 P90996	Human androgen recepto	1.69e+01
32	69	30.7	591	1 W48796	Drosophila melanogaster	2.08e+01
33	69	30.7	732	1 W63715	Rat rsk3 protein.	2.08e+01
34	69	30.7	2441	1 W40058	Cellular transcription	2.08e+01
35	69	30.7	2441	1 R79054	CEB binding protein.	2.08e+01
36	68	30.2	737	1 W18317	Drosophila Deltex prot	2.56e+01
37	68	30.2	737	1 R76640	Deltex protein.	2.56e+01
38	68	30.2	737	1 R76639	Deltex protein.	2.56e+01
39	68	30.2	757	1 R73016	Fish protamine gene tr	2.56e+01
40	68	30.2	1447	1 W81029	Murine pcip protein (o	2.56e+01
41	67	29.8	678	1 R42087	D. melanogaster dorsa	3.15e+01
42	67	29.8	914	1 W24800	Spinocerebellar ataxia	3.15e+01
43	67	29.8	1312	1 W33807	Human ataxin-2.	3.87e+01
44	66	29.3	902	1 R12224	Rat androgen receptor.	3.87e+01
45	66	29.3	1829	1 Y07242	Actin-filament binding	3.87e+01

ALIGNMENTS

RESULT 1
 ID W62841 standard; Protein: 28 AA.
 AC W62841:
 DT 27-OCT-1998 (first entry)
 DE Stenocarpus sinuatus antimicrobial protein.
 KW antimicrobial protein; Infestation; control.
 OS Stenocarpus sinuatus.
 PN W09827805-A1.
 PD 02-JUL-1998.
 PF 22-DEC-1997; AU0874.
 PR 20-DEC-1996; AU-004275.
 PA (REFR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 PI Bower NJ, Goulter KC, Green JL, Manners JM, Marcus JP;
 DR WPI: 98-377279/32.
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
 useful for controlling microbial infestations of plants or mammals
 PS Claim 1; Page 66; 96pp; English.
 CC The sequence is that of an antimicrobial protein which can
 CC be used to control microbial infestations in plants and mammalian
 CC animals.
 SO Sequence 28 AA:
 Query Match 100.0%; Score 225; DB 1; Length 28;
 Best Local Similarity 100.0%; Pred. No. 1.69e-15;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Caps 0;
 Db 1 LDPTRQQLCQMRCCQDEKDPKQOQCK 28
 QY 1 LDPTRQQLCQMRCCQDEKDPKQOQCK 28
 RESULT 2
 ID W62830 standard; Protein: 625 AA.
 AC W62830:
 DT 27-OCT-1998 (first entry)
 DE Macadamia integrifolia antimicrobial protein.
 KW antimicrobial protein; Infestation; control.
 OS Macadamia integrifolia.
 FH key Location/Qualifiers
 FT Peptide 1..28
 FT Protein /note= "signal peptide"
 FT Protein /note= "mature protein"
 PN W09827805-A1.
 PD 02-JUL-1998.
 PF 22-DEC-1997; AU0874.
 PR 20-DEC-1996; AU-004275.

PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR WPI: 98-377279/32.
DR N-PSDB: V42316.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PS Claim 1; Page 43-45; 96pp; English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
SQ Sequence 625 AA;

Query Match 52.9%; Score 119; DB 1; Length 625;
Best Local Similarity 65.4%; Pred. No. 2.93e-04;
Matches 17; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Db 36 DPOTECQCCRCRQESGPRQOYC 61
QY 2 DPTRQQLCOMRCQOQEKDPRQOQOC 27

RESULT 3
ID W62831 standard; Protein; 525 AA.
AC W62831.
DT 27-OCT-1998 (first entry)
DE Theobroma cacao antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS Theobroma cacao.
PN WO9827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997; AU0874.
PR 20-DEC-1996; AU-004275.
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR WPI: 98-377279/32.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PS Claim 1; Page 47-49; 96pp; English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
SQ Sequence 525 AA;

Query Match 52.0%; Score 117; DB 1; Length 525;
Best Local Similarity 60.9%; Pred. No. 4.68e-04;
Matches 14; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Db 83 ROYQCCGRCRQESGPRQOYC 105
QY 5 RQOQLCOMRCQOQEKDPRQOQOC 27

RESULT 4
ID R20181 standard; Protein; 566 AA.
AC R20181.
DT 16-APR-1992 (first entry)
DE Sequence encoded by 67 kD T. cacao protein CDNA.
KW Cocoa; flavour; vicillin; seed storage protein.
OS Theobroma cacao.
PN WO9119801-A.
PD 26-DEC-1991.
PF 07-JUN-1991; G00914.
PR 11-JUN-1990; GB-013016.
PA (MRS) MARS UK LTD.
PI Spencer ME, Hodge R, Deakin EA, Ashton S;
DR WPI: 92-024418/03.
DR N-PSDB: Q20377.
PT Recombinant cocoa proteins - are responsible for flavour in cocoa
PT beans and produced in large quantities using yeast and bacterial
PT expression vectors
PS Claim 4; Fig 2; 59pp; English.
CC The inventors claim a 67 kD and 31 kD T. cacao protein, and
CC fragments, and encoding DNAs. The 47 kD and 31 kD proteins are

CC derived from the 67 kD precursor. T. cacao protein CDNA was
CC detected in a cDNA library prepared from immature cocoa beans RNA
CC using a probe based on the AA sequence of a CNR peptide common to
CC the 47 kD and 31 kD polypeptides. Homology searches revealed close
CC homologues between the 67 kD polypeptide and the vicilins, which are
CC seed storage proteins.
SQ Sequence 566 AA;

Query Match 52.0%; Score 117; DB 1; Length 566;
Best Local Similarity 60.9%; Pred. No. 4.68e-04;
Matches 14; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Db 83 ROYQCCGRCRQESGPRQOYC 105
QY 5 RQOQLCOMRCQOQEKDPRQOQOC 27

RESULT 5
ID W62828 standard; Protein; 666 AA.
AC W62828.
DT 27-OCT-1998 (first entry)
DE Macadamia integrifolia antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS Macadamia integrifolia.
PN Key
FT Peptide
FT Location/Qualifiers
FT 1..28
FT /note= "signal peptide"
FT 29..666
FT /note= "mature protein"
PN WO9827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997; AU0874.
PR 20-DEC-1996; AU-004275.
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR WPI: 98-377279/32.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PS Claim 1; Page 34-36; 96pp; English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
SQ Sequence 666 AA;

Query Match 50.2%; Score 113; DB 1; Length 666;
Best Local Similarity 61.5%; Pred. No. 1.19e-03;
Matches 16; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Db 77 DPOTECQCCRCRQESGPRQOYC 102
QY 2 DPTRQQLCOMRCQOQEKDPRQOQOC 27

RESULT 6
ID W62829 standard; Protein; 666 AA.
AC W62829.
DT 27-OCT-1998 (first entry)
DE Macadamia integrifolia antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS Macadamia integrifolia.
PN Key
FT Peptide
FT Location/Qualifiers
FT 1..28
FT /note= "signal peptide"
FT 29..666
FT /note= "mature protein"
PN WO9827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997; AU0874.
PR 20-DEC-1996; AU-004275.
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR WPI: 98-377279/32.

RESULT	10
ID	W14783.standard; protein; 919 AA.
AC	W14783;
DI	22-JUN-1997 (first entry)
DE	Androgen receptor.
KW	Androgen receptor; acidlab fibroblast growth factor; argef;
KW	antitense; benign prostatic hyperplasia; prostate cancer; therapy

Search completed: Sat May 13 09:23:21 2000
Job time : 8 secs.

Search completed: Sat May 13 09:23:21 2000
Job time : 8 secs.

